DOCUMENT RESUME

ED 322 191 TM 015 374

AUTHOR Stuck, Ivan A.
TITLE Correcting Regression Equations for Restriction of Range: Effects on Veterinary Candidate Selection.
PUB DATE Apr 90
NOTE 29p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Admission Criteria; Cohort Analysis; *College Applicants; College Entrance Examinations; *Equations (Mathematics); Estimation (Mathematics); Grade Point Average; Higher Education; Mathematical Models; *Medical Students; *Predictor Variables; Regression (Statistics); Veterinary Medical Education; Weighted Scores
IDENTIFIERS Michigan State University; *Range Restriction; *Regression Correction

ABSTRACT

Predictor weights estimated by using multiple linear regression are biased when there is restriction in the range (RR) of the dependent variable. Standardized multiple regression yields partial correlations as weights for the predictors, and these can be corrected for range difference between calibration and application samples. However, additional bias occurs in weights due to unreliability in the dependent variable. Five cohorts (four graduation cohorts and one 3-year cohort) of veterinary medicine applicants were studied to assess the effectiveness of adjusting regression weights prior to estimating future graduate grade point average (GGPA). The cohorts were from the College of Veterinary Medicine at Michigan State University (East Lansing), with application years ranging from 1981 to 1985 and graduation years ranging from 1985 to 1989, respectively. Accepted and rejected applicants were included. Three graduation cohorts (1985, 1986, and 1987) were used to select and weight predictors. These weighted predictors were used to estimate future GGPA for the two remaining cohorts (1988 and 1989), which were used for validity testing of the estimated beta weights. The correction appeared to impact the rank order validity of only one cohort, but even this finding was not statistically significant. Creation of a simulated data set also showed larger validity coefficients when RR was corrected; however, the margin of difference was not statistically significant. Two situations were identified where correction for range may be desirable: (1) in cases that are highly homogeneous, and (2) for predictor weights that are highly homogeneous. Six data tables are included. (RLC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED322191

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
IVAN A. STUCK
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Correcting Regression Equations for Restriction of Range: Effects on Veterinary Candidate Selection

Ivan A. Stuck

Michigan State University

BEST COPY AVAILABLE

Running head: CORRECTING REGRESSION EQUATIONS

Paper presented at the annual meeting of the American Educational Research Association, 1990, Boston, Massachusetts

TM015374

Correcting Regression Equations for Restriction of Range: Effects on Veterinary Candidate Selection

Ivan A. Stuck

Michigan State University

ABSTRACT

Predictor weights estimated by using multiple linear regression will be biased where there is restriction in the range of the dependent variable. If standardized multiple regression is used it will yeild partial correlations as weights for the predictors, and these can be corrected for range difference between calibration and application samples. The drawback is additional bias in weights due to unreliability in the dependent variable (raw-score regression would not be affected by measurement error in the dependent variable). Until now the literature has been silent with regards to correcting partial validity coefficients (from multiple linear regression). Five cohorts of veterinary medicine applicants were studied to assess the effectiveness of adjusting regression weights prior to estimating future graduate grade-point average (GGPA). Three cohorts were used to select and weight predictors. These weighted predictors, in turn, were used to estimate future GGPA for the two remaining cohorts. The correction appeared to impact the rank order validity of only one cohort, but even this was not

at the .05 significance level. Creation of a simulated data set also demonstrated larger validity coefficients where restriction of range was corrected, but again, the margin of difference was not statistically significant. Two situations were identified where correction for range may be desirable: (1) where cases are highly homogeneous and (2) where predictor weights are highly homogeneous.

Correcting Regression Equations for Restriction of Range: Effects on Veterinary Candidate Selection

The restriction of range problem occurs when a regression weight estimated for sample A with Criterion Variance A is applied to sample B with a different criterion variance (Criterion Variance B), the resulting criterion estimate will be incorrect. When standardized-score regression is used, the regression weights will be incorrect if criterion or independent variable variance differs from estimation to application samples (see Richards (1982). Formula (1) now exists for the correction of betas (standardized regression weights due to range differences in either or both the predictor and the criterion (Alexander, Carson, Allinger, & Carr, 1987):

$$rho = \left[\frac{rho'^2 - 1}{2rho'} U_x U_y\right] +/- \left[\frac{(1 - rho'^2)^2}{4rho'^2} U_x^2 U_y^2 + 1\right]^{1/2} \quad (1)$$

Where,

rho'=range restricted correlation
U_x= the ratio of restricted to unrestricted standard deviations for variable X
U_y= the ratio of restricted to unrestricted standard deviations for variable Y
+/- :corresponds to the sign of rho'

Nevertheless, until now, this formula has not been reported in application to multiple variable prediction. The object of this study is to estimate the impact of adjustment for range differences in an educational selection situation which uses multiple variable estimation.

Sampling

Sampling theory provides justification for drawing inferences from samples under certain conditions. Suppose that a population exists who share some independently acquired mutual attributes and characteristics but who differ on other attributes and characteristics. If samples are drawn in large enough numbers and in a random manner, we are confident that: (1) randomly sampled, independently acquired characteristics of the sample can be inferred to the population as a whole, and (2) randomly sampled, independently acquired characteristics of the sample can be inferred to any other large random sample of that particular population (these follow from the central limit theorem, see Huntsberger and Billingsley, 1973, pp. 131-134).

Ross (1988) cites Kish for classifying samples as (1) experimental, (2) survey, or (3) investigative, based largely on the quality of the sampling. An experiment provides deliberate treatment with control of extraneous variables by randomization or other means. A survey selects randomly from a defined population in which each member has a specific probability of being studied. In the investigation, however, control is the least. Sampling is by convenience with neither randomization nor

probability sampling. The study of admissions data falls under this latter category. Where cases are not sampled randomly, but are selected according to their value on a particular variable (let's say "selected on IQ" [scholastic aptitude]), observed correlations between that selected variable (IQ) and another (say, academic performance) may be lower than would have been the case if the sample had been sampled randomly (an additional instance of the restriction-of-range problem).

If therefore, selection is on the dependent variable, or if the regression uses standardized variables, regression /validity coefficients for selected data will be artificially low (Richards, 1982). This is the case in sampling to produce a multiple regression selection formula, wherein only selected applicant cases include a GGPA/ criterion. Because the validity coefficient reflects the proportion of true variance to error variance:

$$r_{xy} = s^2_T / (s^2_O = s^2_T + s^2_E),$$

a reduction in true variance resulting from selection of a restricted variable range leaves the error variance intact thus reducing the proportion of true variance to error variance. This is an instance of artifactual error because the proportion of error is inflated due to the improper sampling procedures used.

Other error is due to the vagaries of the sampling process. Nuisance variables (Kirk, 1982), confounding variables, and moderator variables (Allen & Yen, 1979) are common labels referring to another factor which may reduce prediction validity

during the sample selection phase of an investigation. Inasmuch as all members of a population will not equally share access to, nor interest in, graduate admissions; certain population traits and characteristics may be over-represented in a non-random sample of applicants. When such unspecified and uncontrolled-for variables affect performance on the dependent variable, an additional source of error is imposed on the investigation.

Two hazards accompanying the use of non-randomly selected admissions data are therefore: (1) restriction-of-range artifacts and (2) confounded variables.

Measurement

Every measurement can be best regarded as an estimate which includes unknown components of two types of error: (1) unsystematic and (2) systematic. Unsystematic error randomly increases and decreases the measurement value which is observed (relative to the true value of the object or process being measured). Given a large number of measurements, however, the positive and negative errors tend to cancel, leaving a mean value that is virtually the true mean for that set of measurements. Systematic error affects the recorded measurement value in a consistent way (such as always mistakenly using a meterstick instead of a yardstick); regardless of the number of measurements taken, the error remains in the computed mean as well as in the individual measures. Nevertheless, if the nature of the systematic error comes to light, individual measurements

and group statistics may be corrected. Measurement error generally refers to the random kind of error, whereas, systematic error in the measurement is an unaccounted for factor which has a nonrandom influence on the observed scores. If the systematic error is due to instrumentation or procedures, the factor may be called an "artifactual factor". Otherwise, the systematic factors will be attributed to uncontrolled variables in the real data. Of course, errors also differ in level (or magnitude); error that doesn't differ in magnitude across samples is identically distributed error (unilevel), whereas error that does differ across samples (is multilevel) is known as moderated error.

In the context of a distribution, level of error is known as error variance. If error variance is multilevel (or heterogeneous; variance differs across factor levels), it is said to be moderated by that factor. Where error variance is multilevel and, in fact, correlated with the levels of the factor, the error is systematic-- a special case of moderated error.

Where it can be determined that error has systematic or random qualities, the possibility of controlling the error becomes more feasible.

The Reliability of Validity Coefficients

The effect of error on correlation coefficients is more complex than is its effect on observed scores. With no error,

the bivariate correlation is a consistent maximum: the coefficient of the two latent traits. To use a biological analogue, error may be likened to a parasite that invades a "host" variable. From a maximum, latent-trait correlation value, coefficients decline in value as greater levels of error affect the observed scores. This is always true for random error and is virtually always true for systematic error (the exceptional cases being (1) where error adds a constant value to its host variable, or (2) where systematic error is perfectly correlated with its host variable).

Where the error in the observed scores is random, or where the error is systematic relative to its host variable, an unbiased estimate of the expected population coefficient can be computed. It can be computed with precision, moreover, if from a large sample; of course, the resulting correlation coefficient will be attenuated from the latent trait coefficient value due to the random error. Where error varies systematically relative to external influences, however, the computed estimate of the expected population coefficient may be inaccurate in some consistent fashion (biased). The biased estimate of the correlation of the latent traits would, therefore, require a correction of the observed-score correlation.

It is common for error to have attributes of both systematic and random error. It may appear to be normally distributed as in the case of random error, yet also prove to be reducible by the addition of variance controls. This would be the case of a

moderating variable (such as year or location) where error may vary across units of the variable (e.g. times or sites) in either a systematic or a random fashion. By blocking on potential moderator variables (see Neter, Wasserman, & Kutner, 1985) or by using other statistical controls, the moderated error can be removed or reduced.

The restriction-of-range problem is analogous to the problem of unreliability. Both can be accounted for in terms of the proportion of true variance to error variance. Both unreliability and range restriction are reflected in coefficients which are reduced when the proportion of error variance increases. By reducing the range of the variation in the sample of scores (as a consequence of selecting candidates via a cut-score criterion), the proportion of true variance is decreased, and, as a consequence, the proportion of error variance is directly increased (even with no change in the amount of error variance). If the error variance is substantially eliminated, the coefficient approaches the latent-trait value (in the case of parallel tests, that value should be one, though in the case of latent traits, the value could range between positive one and negative one). Linn and Hastings, (1984, p.166) provide a good discussion of the range restriction issue. Variation in range only biases raw-score regression coefficients when the dependent variable range is subject to variation between the calibration and the application samples (Richards, 1982),

although the precision of this unbiased estimate depends heavily on a large sample size.

Richards (1982) discusses the data characteristics displayed in Table 1, which result in error artifacts under (1) raw-score and (2) standardized regression/ correlation. Most notably, raw-score regression coefficient estimates are unbiased by measurement error in the dependent variable, whereas standardized regression coefficients are unbiased by variation in scale units. Neither type of regression is immune from bias due to measurement error in the independent variable.

Although artifactual error in computed statistics may often be reduced through the use of various correction formulas (e.g. for unreliability or for range restriction), these corrections, nevertheless, are limited by the analyst's ability (1) to identify the affected variables (2) to determine levels of variance or reliability under other circumstances. Furthermore,

Table 1

Regression bias factors based on Richards (1982)

BIAS FACTOR	TYPE OF REGRESSION	
	Raw-score	Standardized
Units of measure differ	XX	
Dispersion of independent variable		XX
Unreliability of dependent variable		XX
Dispersion of dependent variable	XX	XX
Unreliability of independent variable	XX	XX
Norm referent scale	XX	XX
Change in test length	XX	XX
Selection on meditating variable	XX	XX

a "corrected" coefficient cannot be assumed to be completely accurate, and may be expected to be conservative (see Linn, Harnisch, & Dunbar, 1981).

Correction of range restriction for a standardized regression equation requires the correction of each partial coefficient. Correction of the raw-score regression model is difficult because it requires the generation of a constant in addition to the transformation of partial coefficients into "b" weights (raw score coefficients).

Some important implications of this theory for selection need to be considered:

(1) If a sample is not randomly drawn its statistics will, nevertheless, represent its population as a whole if all of its relevant characteristics are invariant from member to member (for example, all Girl Scouts are invariant with respect to gender and relatively invariant with respect to age).

(2) If a sample is not randomly drawn, its statistics will also represent its population if the sample is large enough and if the subjects happen to be representative; automobile drivers are random relative to gender and political party preference; ten cars in a line that are picked from a public parking lot may not reflect population composition accurately, but a few hundred cars picked as a block from a parking lot may represent population composition quite accurately relative to gender or political preference. However, sample correlations are apt to be biased

due to moderated error (non-random samples tend to systematically select certain subgroups).

(3) In practice, artifactual differences can be anticipated. Because many factors may be predictive of a particular kind of human performance, one must assume that (a) people may perform at a similar level even though they differ with respect to particular attributes (abilities on several factors may compensate for deficiencies on other factors); and (b) for a given year or location, non-random pressures (self-selection or other non-random selection) must be expected to favor particular factors/attributes resulting in samples which are systematically different from the population as a whole. For instance, a change in requirements for admission to human medicine programs may affect the rate and quality of applications to veterinary medicine.

Sample subgroups may differ in quality and level of preselection prior to inclusion in the sample, due to either self or institutional selection. Aggressive students may be over-represented due to self-selection, and range widely on required aptitudes while students with high verbal skills may be over-represented due to institutional selection and they may range very little on required aptitudes. The effects of these disparities are (a) to create the appearance of a differential validity of any fixed selection formula for the various applicant subgroupings (e.g. verbal, aggressive) and (b) to create the appearance of a differential selection formula validity across

samples (e.g. applicants of different years or locations may differ in their subgroup structure, see Linn, 1983; and Linn & Hastings, 1984). Nevertheless, the validity difference would be largely artifactual, a consequence of range- restriction due to selection.

Differential validity may actually exist independent of the artifactual manifestation, of course; however, a more insightful conceptualization is to attribute this particular validity problem to model misspecification (i.e., a selection formulation lacking in one or more important variables, such as origin of UGPA).

In short, under the uncontrolled conditions of an investigation-level study such as the analysis of graduate admissions data, the likelihood of moderated/ confounded prediction across years or locations is substantial. The result of moderated prediction may be biased regression and validity coefficients.

Multiple Regression

The usual mathematical procedure used in multiple regression yields what is called the ordinary least squares (OLS) estimate of the criterion (or simply the least squares estimate). This term means that the sum of the squared prediction errors is minimized for the data used (see Neter, Wasserman, & Kutner, 1985). The OLS linear model that is calibrated, however, is truly OLS only with respect to the

specific combination of the calibration predictors and the calibration criterion. The OLS correlation (coefficient= R_{OLS}) is the optimal correlation of predicted and actual scores because of the following: (1) scores are calculated by applying the calibrated linear model back onto the calibration predictors, (2) the formula (the linear model) was specifically developed to predict the same criterion data; the subsequent correlation of "ideal" criterion and "ideal" estimated criterion scores is optimal, (3) the OLS correlation is apt to be inflated, to some extent, due to the chance correlation of error with the criterion; when this chance correlation melts away during the application of the linear model to a new sample, the decrease in the coefficient upon cross- validity measurement is called shrinkage.

Multiple regression assumes that (1) responses on the dependent variable are independent, that (2) the variance is constant across cases, that (3) the errors are normally distributed, that (4) the system being modeled is in a steady-state equilibrium, and that (5) errors are uncorrelated (Kenny, 1979, pp.50, 51).

If all of the predictive factors are perfectly represented by the set of predictors and the criterion, the calibrated model will be optimal (although only a perfect correlation if in a totally determined situation where the latent trait correlation is one). Otherwise, as Deegan (1976) and Pedhazur (1982) explain, the following will be true: (1) where superfluous

factors are included among the predictors (Deegan's overspecified model), unsystematic error will be added to the predicted scores when the model is applied, thus causing an underestimate of the validity coefficient of the prediction scores (attenuation due to unreliability); (2) where some factors are omitted from the set of predictors (Deegan's underspecified model which includes the case of omitted independent factors; Deegan, p. 238), systematic error will be added to the predicted scores (this problem can be overcome only by providing the missing predictor data); (3) when a combination of these two situations exists, the model is said to be misspecified (misspecified models have biased parameter estimates which exhibit an interactive character, Deegan, p.238; obviously, without evidence that all predictive factors are appropriately represented, all practical models must be assumed to be somewhat misspecified); (4) where important predictor variables are highly correlated, multicollinearity is said to exist.

Pedhazur (1982) points out that there is differential use of the term, multicollinearity, but its problematic manifestation is biased predictor weights. This problem results in systematic error being added to predicted scores when the calibrated model is applied to non-calibration data (new data). Much work has been done to develop alternate regression techniques to cope with the problem of multicollinearity. Unfortunately, most of these techniques are helpful only in the most severe circumstances (Huberty & Mourad, 1980; Morris, 1986;

Cattin, 1981). Kenny (1979) lists three characteristics associated with low predictor error of measurement. These characteristics are also associated with reduced problems of multicollinearity: (1) a high reliability, (2) a low regression coefficient, and (3) a low predictor intercorrelation. He also adds that multicollinearity decreases as the number of predictors decreases relative to the number of cases.

If, therefore, certain data characteristics exist, the regression procedure will yield an equation which will specify an efficient means of weighting several variables in order to estimate a criterion. Regression equations may suffer from either too few or too many predictor variables. A lack of predictors results in biased regression coefficient estimates, while too many predictors make the coefficient estimates less reliable. High predictor variable intercorrelation may also bias coefficient estimates although this problem is less severe for predictors with high reliability and/or a moderate to low correlation with the criterion.

Method

Sample

The non-random sample consisted of five cohorts; four graduation cohorts and an additional three-year cohort, all from the College of Veterinary Medicine at Michigan State University. The cohorts included both accepted and rejected applicants. Applicants who obtained a vet school GGPA were considered to be

GRADS (this includes the three-year cohort); all others were considered to be non-graduates. Student attrition was estimated to be at less than two students per cohort. Some overlap existed among the cohorts, as some cases were repeat-applicants. While the annual number of applicants changed notably over the years, the number of candidates in-program was relatively constant (see Table 2). All applicants in this sample received scores on the New Medical College Admissions Test (MCAT) and were ranked for admissibility by a selection formula decided upon by the Veterinary College's admission committee. The high admissibility of an applicant could often mean that such an applicant might accept another candidacy elsewhere, therefore, some non-graduates were of this caliber.

Table 2

The study sample: Five cohorts of veterinary applicants

APPLICATION YEAR	GRADUATION YEAR	NUMBER OF APPLICANTS	NUMBER OF GRADUATES
1981	1985	327	90
1982	1986	327	89
1983	1987	310	103
1984	1988	273	88
1985	1989	248	101

Three graduation cohorts (1985, 1986, 1987) were used in the calibration sample for estimating regression weights and the other two cohorts (1988, 1989) were used for validity testing of the estimated beta weights.

Predictors

Predictor variables were the following: CUMGPA (cumulative UGPA), PVUGPA (veterinary prerequisites UGPA), honor points, prerequisite course honor points, credits without pass/fail credits, average credits per term, summer credits, total credits, number of terms, pass/fail credits, the Medical College Admissions test subtest scores, age, sex, interview scores I and II, work experience, veterinary experience, and narrative work sample.

Predictors were available from admissions data variables. Because many of these variables were not of interval scale, and because some of them were composites of several variables, such variables were deleted. Composites were sacrificed for single variables. The twenty-two variables are listed in Table 3.

Criterion

Program graduate grade point average (GGPA) served as the criterion.

Analysis

Stepwise multiple regression was used to select and weight predictors, using data years 1981 to 1983. The SPSSx default settings were allowed to control admission and exclusion of predictors from the equation (significance level= .05). The resulting regression beta weights for the selected variables were retained in (1) their restricted form but were also

Table 3

Predictors used and levels of intercorrelation

PREDICTORS	INTERCORRELATION LEVELS >.45	<.45
cumulative UGPA	XX	XX
veterinary prerequisites UGPA	XX	XX
veterinary prerequisites points	XX	
honor points	XX	XX
credits	XX	
average term credits	XX	XX
sum of pre-veterinary credits	XX	XX
total credits	XX	
number of terms	XX	
pass/fail credits	XX	XX
MCAT Biology	XX	
MCAT Chemistry	XX	XX
MCAT Physics	XX	
MCAT Quantitative	XX	
MCAT Reading	XX	XX
MCAT Science	XX	
age	XX	
sex	XX	XX
veterinary experience	XX	
work experience	XX	XX
activities and achievements rating	XX	XX
narrative writing sample rating	XX	XX

(2) corrected using the Alexander et al. (1987) correction formula (equation 1). Corrected and uncorrected regression weights were applied to 1984 and 1985 admissions data to estimate GGPAs. The stepwise multiple regression using pooled datasets from three cohorts selected only two variables: (1) UGPA (undergraduate GPA) and (2) the Science Reasoning subtest of the New MCAT. Their beta weights were .466 and .218 respectively. Cases with both an authentic and an estimated GGPA were selected for the test of estimation methods. The estimated GGPAs were

ranked (each for 1984 and 1985), as likewise were the authentic GGPAs. Six sets of rankings were thus produced:

(1984) a. authentic GGPA ranks
b. uncorrected GGPA estimate ranks
c. corrected GGPA estimate ranks

(1985) a. authentic GGPA ranks
b. uncorrected GGPA estimate ranks
c. corrected GGPA estimate ranks.

For each year, each of two estimated score rankings were correlated with the authentic GGPA ranking, resulting in two tests of corrected vs uncorrected regression coefficients.

Results

Table 4 displays the rank order and ordinary correlations for corrected and uncorrected GGPA estimates for 1984 and 1985 datasets. The corrected coefficients seem to result in better prediction for 1985 data, whereas the uncorrected coefficients appear more predictive for 1984. Using the Fisher r to Z transformation (Hays, 1973) to test for correlation differences between the 1985 coefficients for the corrected and uncorrected GGPA estimates, the one-tailed test indicates a .09 probability level, short of the conventional .05 standard of significance. The set of 1984 coefficients are even less disparate and hence (obviously) also non-significant.

Table 4
Correlations: GGPA by estimated GGPA (corrected and uncorrected)

TYPE AND YEAR OF COEFFICIENT	COEFFICIENT:GGPA X ESTIMATE
1984 Rank Order Uncorrected	.60
1984 Rank Order Corrected	.56
1985 Rank Order Uncorrected	.36
1985 Rank Order Corrected	.47
1984 Pearson Uncorrected	.57
1984 Pearson Corrected	.56
1985 Pearson Uncorrected	.36
1985 Pearson Corrected	.48

The Pearson and rank order coefficients are similar in magnitude. The rank order coefficients are, nevertheless, more appropriate for comparing estimates in a selection context because selection is more likely to be based on rank than on absolute value. In addition, it is possible for differences to exist between Pearson coefficients despite no differences occurring between rank order correlations. For this data, however, the coefficients are virtually identical.

Discussion

This study is a preliminary effort to understand the impact of the application of a correction for attenuation on a regression equation used for educational selection. Both the adjustments for coefficient attenuation, (1) due to range

differences and (2) due to reliability differences, are necessary procedures for providing accurate regression equations, yet there seems to be no literature confirming the validity of such adjustments on regression partial coefficients.

The present study observed only two replications of a test to compare the precision between score estimates using range corrected and uncorrected regression weights. Even had there been findings of significant differences for these two replications of this test of coefficients (using real data), empirical evidence of the need for coefficient correction would have remained inconclusive; two cross validation samples may have been well predicted by the corrected coefficients merely by chance. Particularly for tests of correlation coefficients, size of N (or sampling error) is a crucial consideration (see Hunter, Schmidt, and Jackson, 1983). Short of the ideal massive data base, a more liberal level of significance is appropriate for this type of test. Had a significance level of .10 been specified for this comparison, one hypothesized outcome would have been statistically significant. Unfortunately, even such an event would not compell us to dismiss the explanation of sampling error as the cause of the size of the coefficient difference. Ultimately, the issue of coefficient correction probably needs to rest on the rational basis of the necessary consequences of (1) range variation and of (2) measurement error on regression weight estimates.

Few would likely debate the wisdom of adjusting a dominant predictor for range differences, however, the legitimacy of adjusting lesser predictors is less obvious. One important consideration has been overlooked, however, which should reduce concerns about the role of lesser predictors in the estimation of variable weights. Because the object of selection is (usually) to select the most meritorious or otherwise appropriate candidates, relative as opposed to absolute prediction is a sufficient goal. It happens to be true that estimation weights can vary considerably from precision in an absolute sense without impairing the relative order of the estimates; it requires a fairly dramatic deviation from the true variable weights to reduce the validity of the resulting estimated rank order. However, if relative estimation is insensitive to weighting errors, then, why be concerned about adjusting them for restriction in range? There are two reasons that can be offered: rank order does become more vulnerable to error as (1) the homogeneity of the sample increases, and (2) as the variables representing the dominant prediction factors display homogeneous regression weights.

In the first instance, rank order validity is more susceptible to attenuation as cases increase in homogeneity. In the second instance, changes in regression weights only become important for estimated rankings when the "wrong" predictors dominate (a potential consequence of range restriction and of unreliability) in the stepwise multiple regression.

Table 5 presents two examples of the effects of range restriction for a simulated data set of 1000 cases. The criterion is estimated by two variables which are very similar in predictive validity (.5 and .47). When the stronger predictor is used to select only cases above the predictor mean, the (1) regression weights estimated from the selected sample and (2) the same weights after correction for range difference are virtually as valid for ranking ability for the 1000 cases (.7856 vs .7938). After selecting the top 47 cases on the dominant predictor, the disparity of rank validity increases (.6896 vs .7449).

Table 5

Range restriction effects for two variables with similar weight

	Correlation of Ranks (PPM) Restricted Weight Estimates	Correlation of Ranks (PPM) Corrected Weight Estimates
Cut-score at Mean n= 492	.786 (.706)	.794 (.712)
Cut-score 1.25 above Mean n= 47	.690 (.629)	.745 (.674)

N= 1000

In Table 6, the simulated dataset includes a third variable which has low intercorrelation with the other two predictors (in Table 5). The data simulates three sets of equivalent forms scores for the same set of abilities. Restricted and corrected coefficients are computed from cases selected at the same high

Table 6

Range restriction effects for three equivalent forms

	Correlation of Ranks (PPM) Restricted Weight Estimates	Correlation of Ranks (PPM) Corrected Weight Estimates
1. Cut-score 1.25 above Mean n= 37	.750 (.665)	.764 (.678)
2. Cut-score 1.25 above Mean n= 39	.788 (.701)	.813 (.729)
3. Cut-score 1.25 above Mean n= 45	.726 (.670)	.745 (.686)

N= 1000

value of the dominant predictor as in the highly selected example in Table 5. The three sets of corrected coefficients appear to yeild more valid rankings (.76 vs .75, 81 vs 79, and .75 vs .73) although all coefficients are statistically equivalent (p= .05).

The real and simulated data findings from this study suggest that only small gains in selection validity are likely from correcting regression coefficients for restriction in range. This small validity increment may be more important for selection "fairness" than for its practical benefit to the selection process as a whole.

Future study should simulate (1) selection at both the upper and lower extremes of the distribution and (2) selection on one predictive factor variable of a multifactor criterion. Each of

these two types of situations could legitimately represent real data relations. Because these situations can involve a greater restriction effect, these situations have the potential of reflecting a more dramatic correction effect than the present restriction simulation.

REFERENCES

Allen,M.J.,Yen,W.M. (1979) Introduction to Measurement Theory. Monterey, California: Brooks/Cole.

Alexander,R.A.,Carson,K.P.,Alliger,G.M.,Carr,L. (1987) Correcting doubly truncated correlations: An improved approximation for correcting the bivariate normal correlation when truncation has occurred on both variables. Educational and Psychological Measurement, 47, 309-315.

Cattin,P. (1981) The predictive power of ridge regression: Some quasi-simulation results. Journal of Applied Psychology 66:3, pp.282-290.

Deegan,J.Jr. (1976) The consequences of model misspecification in regression analysis. Multivariate Behavioral Research, April, 1976, pp.237-248.

Hays,W.L. (1973) Statistics for the social sciences (2nd ed.). Chicago: Holt, Rinehart and Winston, pp. 664-667.

Huberty,C.J.,Mourad,S.A. (1980) Estimation in multiple correlation / prediction. Educational and Psychological Measurement, 40:1, pp.101-112, Spring.

Hunter,J.E.,Schmidt,F.L.,and Jackson,G.B. (1983) Meta-analysis: Cumulating research findings across studies. Beverly Hills: Sage, pp. 92-94.

Huntsberger,D.V.,Billingsley,P. (1973) Elements of Statistical Inference. Boston: Allyn and Bacon, pp.131-134.

Kenny,D.A. (1979) Correlation and causality. Wiley & Sons, New York.

Kirk,R.E. (1982) Experimental Design (2nd ed). Monterey, California: Brooks/Cole.

Linn,R.L.,Harnisch,D.L.,Dunbar,S.B. (1981) Corrections for range restriction: An empirical investigation of conditions resulting in conservative corrections. Journal of Applied Psychology, 66:6, pp.655-663.

Linn,R.L. (1983) Pearson selection formulas: Implications for studies of predictive bias and estimates of educational effects in selected samples. Journal of Educational Measurement, 20:1, pp.1-15.

Linn,R.L.,Hastings,C.N. (1984) Group differentiated prediction. Applied Psychological Measurement, 8:2, pp.165-172.

Morris,J.D. (1986) Selecting a predictor weighting method by PRESS. Educational and Psychological Measurement, 46, p.853-869.

Neter,J.,Wasserman,W.,Kutner,M.H. (1985) Applied Linear Statistical Models, (2nd. Ed.) Homewood, Illinois: R.D.Irwin, Inc., p.10.

Pedhazur,E.J. (1982) Multiple regression in behavioral research: Explanation and prediction (2nd. ed.) . Chicago: Holt, Rinehart, and Winston.

Richards,J.M. (1982) Standardized versus unstandardized regression weights. Applied Psychological Measurement, 6(2), 201-212.

Ross,K.N. (1988) Sampling. In: Keeves (ed) Educational Research, Methodology, and Measurement: An International Handbook. Pergamon. New York, pp. 527-537.